DOCUMENT RESUME

ED 290 174 CS 506 003

AUTHOR Dawson, Pamela J.; Spitzberg, Brian H.
TITLE Improving Communicative Competence: Validation of a Social Skills Training Workshop.
PUB DATE Nov 87
NOTE 32p.; Paper presented at the Annual Meeting of the Speech Communication Association (73rd, Boston, MA, November 5-8, 1987).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Communication Research; *Communication Skills; Higher Education; *Interpersonal Communication; *Interpersonal Competence; *Program Effectiveness; Remedial Programs; Speech Communication; Test Validity; Training Methods; *Workshops
IDENTIFIERS Conversational Skills Rating Scale; Simulated Social Interaction Test

ABSTRACT

The effectiveness of a social skills training workshop was assessed by comparing the rated competence of participants in an Interpersonal Skills Training Program to the rated competence of nonparticipants. Subjects' self-ratings were included. This comparison was operationalized through a pretest-posttest design with 12 experimental and 22 control subjects. The assessment instruments consisted of Spitzberg and Hurt's (1987) Conversational Skills Rating Scale (CSRS) and Curran's (1982) Simulated Social Interaction Test (SSIT). Two rating judges were utilized. Results, although modest, are in the expected direction. Assessment of competence with the CSRS failed to reveal any significant improvement in the experimental group relative to controls. However, the SSIT did reveal significant improvement of the rated skills and anxiety of experimental subjects while the control group showed no significant improvement. In addition to assessing the effectiveness of the training program, this study sought to establish further validity support for the CSRS. As expected, the CSRS revealed a positive correlation to SSIT ratings. (Author)

**********
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
**********

NC

ED290174

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Pamela J. Dawson

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

# IMPROVING COMMUNICATIVE COMPETENCE: VALIDATION OF A SOCIAL SKILLS TRAINING WORKSHOP

Pamela J. Dawson
841 Londonderry Lane, #4
Denton, TX 76205

Dr. Brian H. Spitzberg
North Texas State University
(Soon to be University of North Texas)
P.O. Box 5266
Denton, TX 76203-5266

**Abstract**

The effectiveness of a social skills training workshop was assessed by comparing the rated competence of participants in an Interpersonal Skills Training Program to the rated competence of nonparticipants. Subjects' self-ratings were included. This comparison was operationalized through a pretest-posttest design with 12 experimental and 22 control subjects. The assessment instruments consisted of Spitzberg and Hurt's (1987) Conversational Skills Rating Scale (CSRS) and Curran's (1982) Simulated Social Interaction Test (SSIT). Two rating judges were utilized. Results, although modest, are in the expected direction. Assessment of competence with the CSRS failed to reveal any significant improvement in the experimental group relative to controls. However, the SSIT did reveal significant improvement of the rated skill and anxiety of experimental subjects while the control group showed no significant improvement. In addition to assessing the effectiveness of the training program, this study sought to establish further validity support for the CSRS. As expected, the CSRS revealed a positive correlation to SSIT ratings.

Paper presented at the Speech Communication Association Conference, Boston, MA, November 1987

CS506003



BEST COPY AVAILABLE



The plethora of training programs, measurement methodologies, and conceptualizations in the area of social and interpersonal skills can be traced to a core problem: the lack of a unifying, integrative framework. Clinicians have little basis for selecting, organizing and implementing training programs, given the enormous diversity and breadth of activities available. As yet, no single training approach has proven unequivocally to be superior empirically to alternative approaches (see Curran, 1977, 1979a, 1979b; Hersen & Bellack, 1976). Furthermore, no single diagnostic technique has proven to be superior to other assessment techniques (Bellack, 1979, 1983; Eisler, 1976; Hersen & Bellack, 1977; Liberman, 1982; Spitzberg, in press). These are not unrelated trends. Validation of training methodologies presupposes valid measurement techniques to ascertain training effectiveness. Furthermore, until there are clear conceptual models of the nature of social skills and the learning process underlying these skills, it will be difficult to develop measures that accurately reflect the skills and abilities being trained. The study reported herein is directed toward these dual purposes: to examine the validity of an interpersonal skills measure, and to assess the effectiveness of an integrative model of interpersonal competence and a training program designed from its perspective.

Although social skills training has been a topic for research for decades, there has been little agreement on the most effective model for the design and implementation of training programs. Different grounding theories of learning lead to very different types of skills training programs (Ellis & Whittington, 1981).

Recently, training programs with eclectic learning approaches have increased in popularity (Kurtz, Marshall & Banspach, 1985). These social skills training programs are combining role-playing, modeling, cognitive restructuring, and behavioral conditioning with successful results, at least for short-term changes in behavior (Ellis & Whittington, 1981). However, while the eclecticism in social skills training shows promise for the successful transfer of learning, there is a distressing lack of conceptual coherence upon which to build and validate social skills training programs.

Social skills training programs have generally focused on social skills deficits; that is, extinguishing ineffective behaviors. Some training programs have addressed shyness or heterosexual-social anxiety (see Curran, 1977). Other training programs have achieved wide popularity in a specific remedial area, e.g., assertiveness training (see Galassi & Galassi, 1976). Still other training has focused on improving psychiatric patients' and mentally retarded individuals' abilities to function more effectively in society and to be more self-reliant (e.g., learning to make requests and to ask questions).

Regardless of the success of the various social skills training programs, the concept of social skills is rarely given adequate attention in research studies. Spitzberg and Cupach (1984) offer a comprehensive conceptualization of social skills which they model in terms of relational competence. Because competence is a quality perceived by others (McFall, 1982), it is a relational concept:

> Relational competence can be defined conceptually as the extent to which objectives functionally related to communication are fulfilled through cooperative interaction appropriate to the interpersonal





> context. Therefore, relationally competent communication is conceptualized as a function of perceived appropriateness and effectiveness (Spitzberg & Cupach, 1984, p. 100).

This model of relational competence offers a potential solution to the lack of conceptual coherence in the research studies of social skills training. Their model of relational competence is organized around three components: motivation, knowledge, and skill. If eclectic learning activities can be organized according to the three components of relational competence and interpersonal skills are positively affected, a major step will have been taken in validating a unifying framework for social skills training.

The effectiveness of training programs based on a comprehensive model of communication competence has not been tested. This study tests the effectiveness of a social skills workshop, or interpersonal skills training program (ISTP), based on a comprehensive model of relational competence. This purpose is achieved by examining the effects of the training program on the ratings of participants' interpersonal competence as determined by a pre- and postassessment, compared to a control group of training program nonparticipants, and to self-reports.

Unlike many training techniques, the ISTP takes a more comprehensive view of the nature of interpersonal competence. For example, many training programs view social skills problems primarily as motivational deficits (e.g., rhetoritherapy, systematic desensitization, cognitive restructuring) or skills deficits (e.g., heterosocial skills, assertiveness skills). The relational competence model assumes that communication difficulties may have multiple etiologies. Persons may experience communication failure

because of positive (i.e., they may not recognize or value the contextual goals) or negative (i.e., they are incapacitated by anxiety or fear) motivational states, they may lack the substantive or procedural knowledge required, or they may simply lack the behavioral skills needed to perform competently. While not entirely consistent, the results of this study lend support to the comprehensive framework embraced by the ISTP. Indeed, the ISTP does not represent an innovation in training content per se, but a new way of organizing and making sense of existing eclectic training techniques under the aegis of a single program.

The primary thrust of this study is to examine the efficacy of an interpersonal skills training program (ISTP). In order to address this issue, the validity of the measures used for assessment must be addressed. If the Conversational Skills Rating Scale (CSRS) and Simulated Social Interaction Test (SSIT) are valid measures of interpersonal skill, they should be positively correlated to each other.

HYPOTHESIS 1:

The Conversational Skills Rating Scale (CSRS) is positively related to the Simulated Social Interaction Test (SSIT).

Presuming that there is validity for the CSRS and SSIT, the expectation follows that subjects receiving training in interpersonal skills will show significant and positive improvement relative to subjects receiving no training.

HYPOTHESIS 2:

Participants in an interpersonal skills training program will show significant positive improvement in their rated interpersonal skills, while a control group of subjects not participating in the training





program will show no significant change in their rated interpersonal skills.

Method

At a large southwestern public university, an Interpersonal Skills Training Program (ISTP) has been developed based on Spitzberg and Cupach's model of communication competence. It is conducted each semester for students who receive low ratings on their in-class interpersonal competence activity by their communication instructor. One of those training programs provided an opportunity to test the effectiveness of the ISTP, and by implication, its eclectic training model on the participants' resultant rated competence.

In order to determine what improvement, if any, such a training program could facilitate, a pretest and posttest design with an experimental group of participants and a control group of nonparticipants was chosen. Such a design would provide relatively unambiguous results with clear comparisons of posttest ratings of competence to the pretest ratings, as well as a comparison of the experimental group to a control group. Isolation of observed differences would therefore be enhanced.

Selection of subjects

Study subjects were recruited from 38 sections of the basic communication course. Students who qualified for training program participation but could not attend served as control subjects.

All training program participants (experimental group) and nontreatment volunteers (control group) were videotaped in dyadic conversations with each other and in simulated role-plays individually. Videotaping of subjects was done the weak preceding and the second week

following the training program; a period of three weeks separated the pretest session and the posttest session. Conversation pairs consisted of a variety of experimental and control subject combinations: experimental/control, experimental/experimental, and control/control. Matching was based entirely upon the subjects' signing up for pretest and posttest time slots consistent with their individual schedules.

After all subjects were videotaped chronologically, rating tapes were re-recorded to mix the pretest and posttest conditions, consolidating conversations for CSRS rating purposes and consolidating role-plays for SSIT rating purposes.

Assessment instruments

In his analysis of social skills and interpersonal competence measures, Spitzberg (1986a, 1986b) has identified as many as 138 relevant measures. However, despite the number of measures, Spitzberg concludes that few of them have been extensively researched. Such diversity of assessments makes it clear that there is a need for a flexible, convenient, and validated measure of skills. For the assessment of the social skills or communication competence of videotaped subjects, the ideal measurement should be simple and uncomplicated yet valid and reliable enough to assess communication competence accurately. The Conversational Skills Rating Scale (CSRS) (Spitzberg & Hurt, 1987) was developed to meet this need.

CSRS. Developed initially through pilot studies and literature search, molecular-level, discrete behaviors that can be observed and rated (Spitzberg, 1985) comprise 25 of the 30 items on the CSRS. The skill items identified on the CSRS include such behaviors as "use of eye contact;"



BEST COPY AVAILABLE



"vocal volume (neither too loud or too soft);" "asking of questions." The remaining five items are molar-level evaluations of overall conversational skill, expressiveness, altercentrism, composure, and appropriateness/effectiveness. The 25 molecular items are rated on a 5-point Likert-type scale, anchored as follows: 1 = INADEQUATE (use was awkward, disruptive, or resulted in a negative impression of communicative skills); 3 = ADEQUATE (use was sufficient but neither very noticeable nor excellent. Produced neither positive nor negative impression); 5 = EXCELLENT (use was smooth, controlled, and resulted in positive impression of communicative skills).

In addition to providing diagnostically useful assessment of 25 molecular conversation behaviors and 5 molar competence ratings, the CSRS is self-explanatory and easy to use by untrained raters.

Spitzberg and Hurt (1987) assessed the CSRS both as a self-report measure and as an observational measure of in-class student "get-acquainted" conversations. The results of that study supported the reliability, convenience and powerful relationship of the behavioral items to the molar ratings of interactants' communicative performance.

SSIT. In addition to ratings of 25 molecular behaviors provided by the CSRS, another measure was desired to serve as a comparison for the CSRS. Curran's (1982) Simulated Social Interaction Test (SSIT), a proven role-play instrument, was chosen because of its empirical validity and for its ease of administering. Careful training can ensure high interrater reliability for judges.

The SSIT is composed of a role-play orientation narrative and a script of eight problematic situations that are presented to individual subjects

who respond verbally in role-play fashion. The eight role-play situations consist of three parts: 1) the narrator describing the situation; 2) a confederate (or narrator) delivering the situational prompt; and 3) the subject's response to the prompt. The eight situations address situations many people have trouble with, e.g., dealing with disapproval or criticism, expressing interpersonal warmth, and receiving complements. For example, the situation dealing with disapproval or criticism is set up for the respondent as follows:

NARRATOR: You are a work, and one of your bosses has just finished inspecting one of the jobs that you have completed. He says to you:

CONFEDERATE: "That's a pretty sloppy job. I think you could have done better."

The subject then responds to the confederate's remark as if he or she were actually in that situation, e.g., "I'm sorry. In what way did I ...." Raters make molar evaluations on an 11-point Likert-type scale of the subject's response to each situation on two dimensions: social skillfulness and anxiety.

The SSIT has been tested extensively (see Curran, 1982) and, compared to other measures of competence, the SSIT is one of the best validated measures available (Spitzberg, 1986b).

Raters

Because time requirements for rating subjects would be demanding (20 hours), there were two volunteers from the cadre of graduate students in communication. If interrater reliability were controlled, two raters could provide the necessary ratings. Because the volunteer raters were graduate assistant instructors in the basic communication course, they were versed





in using the CSRS instrument in rating students in their classes; training was thus streamlined and compressed to address the SSIT alone.

One training session was conducted that 1) familiarized the raters with the SSIT role-play situations, 2) provided criterion behaviors as behavioral observation anchors, 3) utilized videotaped sample SSIT role-plays as practice, and 4) included discussion and comparison of rater practice ratings. After the first round of practice ratings, criterion behaviors were again discussed in order to identify perceptual discrepancies and to anchor raters' perceptions of subjects' skillfulness and anxiety behavior.

Two more rounds of practice ratings followed by discussion of behavioral anchors and comparison of ratings were conducted; rater agreement improved with each round, the third round producing acceptable agreement levels. Percentage of near-agreement of the SSIT items on the first two practice ratings averaged about 63%, improving to 81% agreement on the third round. Considering the similarity of the two raters' graduate and teaching experiences, and the positive outcome of the intensive training session, rater agreement ranging between 70% and 80% on the ratings of the study subjects was expected.

## Results

The average construct scores are examined first; that is, the ratings by the two raters on the major variables are summed and divided by two to provide a mean construct rating. In cases where these findings do not meet with expectations, or are anomalous, individual rater analyses are performed. It is hoped that by examining the individual raters in such cases the reasons for the anomalous results can be elucidated.

Reliability

Interrater reliability was assessed by Pearson correlations between raters' scores on the CSRS, SSIT/skill and SSIT/anxiety constructs. Coefficients, shown in Table 1, are low in spite of preliminary rater training and indications in prior research that the SSIT should have achieved higher coefficients (Curran, 1982).

---

TABLE 1 ABOUT HERE

---

Despite these discouraging results, there are several reasons to continue analysis. First, published research has reported lower reliabilities for subjective rating instruments (e.g., Waltz & Gough, 1984). Second, for exploratory purposes, results should still be examined to determine if there is reason for further research. Third, to the extent that significant results are found despite rater disparities, we can conclude that the constructs studied are powerful enough to overcome these statistical problems. And last, given that validity is ultimately a more important question than reliability, experts have recommended using averaged scores to enhance both the reliability and validity of ratings (Strahan, 1980; Horowitz, Inouye & Siegelman, 1979).

Internal consistency of the measure was assessed by the coefficient alpha reliability. Since this statistic is a function of sample size, the coefficient produced can be considered an extremely conservative estimate of internal reliability. The CSRS produced coefficients for the pretest and posttest conditions for Rater 1 of .75 and .85, respectively. The





higher coefficient for the posttest condition suggests a possible learning effect in using the SSIT.

Examination of the coefficient alpha for the SSIT is more complicated. In addition to having a small sample broken down on the subscales, the SSIT constructs of skill and anxiety consist of only eight items each. Since coefficient alpha is a function of the number of items and sample size, the coefficients produced are certainly deflated. Table 2 displays the coefficient alpha for the SSIT skill and anxiety constructs broken down by rater and condition. Interestingly, opposite learning effects seem to have occurred by the raters. Rater 1 appears to have become less reliable whereas Rater 2 became more reliable.

---

TABLE 2 ABOUT HERE

---

Validity

Support was found for Hypothesis 1; the CSRS instrument correlated positively to the SSIT instrument ($r$=.55, $p$< .01). It is interesting to note that, in addition to the extensive research literatures supporting both the CSRS and SSIT, the averaged CSRS competence pretest ratings were significantly related to SSIT/skill posttest ratings three weeks later ($r$=.55, $p$< .001) and to SSIT/anxiety pretest ratings ($r$=.67, $p$< .001). This provides evidence of the utility of the CSRS instrument.

As would be expected, SSIT/skill pretest ratings are significantly related to SSIT/anxiety pretest ratings ($r$=.74, $p$< .001), SSIT/skill posttest ratings ($r$=.45, $p$< .01), and SSIT/anxiety posttest ratings ($r$=.48,

$p < .001$). CSRS competence pretest ratings are substantially related to CSRS competence posttest ratings ($r = .78$, $p < .001$).

In short, the constructs appear to be relating to one another in ways that would be expected and, in some cases, reveal impressive power for such a small sample.

Test of the Workshop

The essential purpose of this study was to assess the effectiveness of an interpersonal skills training program (ISTP). To determine this, three constructs were used as dependent variables: CSRS ratings, SSIT/skill ratings and SSIT/anxiety ratings. As discussed earlier, these constructs are averaged across raters except in instances where results are counterintuitive. The expectation for these constructs was that each would show significant increases in the experimental (ISTP) condition but not in the control condition.

---

TABLE 3 ABOUT HERE

---

Hypothesis 2 was partially supported. CSRS ratings shown in Table 3, did not reveal a significant change in the experimental or control conditions. The female experimental subjects show some improvement over the controls, while male experimental subjects do not. Figures 1 and 2 illustrate this disparity.

---

FIGURE 1 ABOUT HERE

---





---

FIGURE 2 ABOUT HERE

---

In order to assess the reason for observing no significant change in the experimental condition, an analysis of rater differences in the CSRS was performed (Table 4). Raters 1 and 2 varied significantly in their ratings of the experimental group. Rater 1 found no significant difference in those subjects' behavior in a pretest/posttest comparison. On the other hand, Rater 2 did indeed perceive significant improvement in the experimental subjects' behavior in the CSRS pretest/posttest comparison. In short, it appears that Rater 1 perceived no such change, as assessed by the CSRS instrument. A possible reason for this is addressed in the discussion below.

---

TABLE 4 ABOUT HERE

---

In the examination of the SSIT/skill ratings, results (shown in Table 5) were found that supported the efficacy of the skill training program. Subjects were rated as significantly more skillful in the experimental condition while no change was observed in the control condition.

---

TABLE 5 ABOUT HERE

---

The SSIT/anxiety ratings in Table 6 reveal a similar pattern to the SSIT/skill ratings. Subjects in the experimental condition were perceived

as significantly less anxious after the training program, while no change was observed in the control subjects.

---

TABLE 6 ABOUT HERE

---

Analysis of variance using the SSIT/skill posttest ratings as the dependent variable crossed by condition and sex reveals an interaction effect. While condition produced a nonsignificant main effect and sex a significant main effect, the variables reveal a significant interaction effect that explains almost 22% of the variance ($r^2$=.46) (see Table 7). Figures 3 and 4 illustrate comparisons of SSIT/skill mean ratings by treatment condition and sex of subjects.

---

TABLE 7 ABOUT HERE

---

---

FIGURE 3 ABOUT HERE

---

---

FIGURE 4 ABOUT HERE

---

Using the SSIT/anxiety posttest ratings as the dependent variable crossed by condition and sex, analysis of variance reveals significant main effects and nonsignificant effects for each of the variables in isolation and interaction. The overall model approaches significance and explains a





substantial amount of variance ($r^2$=.21) (see Table 8). Figures 5 and 6 illustrate comparisons of SSIT/anxiety mean ratings by treatment condition and sex of subjects.

---

TABLE 8 ABOUT HERE

---

---

FIGURE 5 ABOUT HERE

---

---

FIGURE 6 ABOUT HERE

---

## Discussion

The purpose of this study was to determine if an interpersonal skills training program (ISTP) would facilitate significant improvement in participants' interpersonal skills or competence. For the most part, the results have proved that this training program has indeed shown effectiveness, despite some reliability problems.

### Rater Differences

Although reliability coefficients were far below the .70 level anticipated for this study, there is encouragement that these low coefficients can be improved substantially with more rigorous rater training.

In contrast to the rater training in this study, Curran (1982) trained lay people as well as communication specialists were utilized as rater

trainees in a very lengthy and thorough procedure that included extensive explanation of criterion behaviors rationale and 12 hours of practice ratings by rater trainees. Only trainees with a degree of agreement with the criterion ratings of .80 or better were selected to rate the experimental tapes. These training procedures, proctored by a research team, lead to high interrater reliability and agreement with criterion ratings.

Unfortunately, such a pool of volunteer (or paid) raters was not practical for this study due to the extreme time demands of such participation on graduate student volunteer raters and to budgetary limitations. With these constraints, one 5-hour training session was designed to anchor behavioral criteria for the SSIT ratings, and interrater reliability was close to .70 at the end of that training session.

Because both volunteer raters were also teaching assistants in the communication department, they were already familiar with the CSRS rating instrument. Extensive rater training of the CSRS has been an integral component of teaching assistants' training each fall. Previous studies validating the CSRS instrument (Spitzberg & Hurt, 1987; Spitzberg, 1985; Spitzberg, 1986a; Spitzberg, 1986b) have found generally high interrater reliability, localizing rater differences.

Despite ratings of significant improvement in the experimental group as compared to the control group on the SSIT by both raters, the CSRS ratings proved contradictory to expectations: Rater 1 observed no significant improvement in either experimental or control group subjects, while Rater 2 did observe significant improvement in the experimental group but not in the control group.





Several factors could be contributing to such disparity. One factor that stands out is that the two raters were perceiving subjects' behavior differently. In fact, some items reveal an inverse relationship: the higher one rater scored the behavior, the lower the other rater scored it. This phenomenon occurred on only a few, isolated items, but it is perplexing nonetheless, considering the raters' common teaching experience.

Another factor that could have contributed to the rater differences is that, although both raters were peer teaching assistants, the raters participated in different teaching assistant training programs (one year apart). The training methods and criterion behaviors utilized in the two different sessions to anchor ratings on the CSRS could very well have been substantially different to cause lower coefficients. It is even possible for each rater to be highly correlated with his own training group, yet not significantly correlated with another teaching assistant from a different training group. It would be interesting to see how well individuals with no specialized knowledge in interpersonal communication might correlate with one another with training similar to that in this study. Curran (1982) had success with such raters.

Finally, the measures used in this study are subjectively scaled instruments, and the criteria for judging competence are inherently subjective. Consequently, high reliability is not expected. Therefore, averaged scores are more consensual in the present study.

Subjects' Self-Ratings

The subjects self-ratings were not correlated with any other ratings except other self-ratings. The CSRS posttest self-ratings were predicted by the pretest self-ratings; the subjects were rating themselves

consistently over time. The more a subject perceived and rated improvement on the SSIT/skill dimension between pretest and posttest, the higher those subjects' CSRS self-ratings were.

Effectiveness of the Training Program

Despite the problems with interrater reliability, the results show improvement of the experimental group subjects in comparison to the control group subjects. Although the overall CSRS competence ratings indicate no significant change in the experimental subjects, means are in the expected direction. The SSIT/skill and SSIT/anxiety ratings are stronger in showing significant change in the experimental group over the control group.

The results show movement in the expected direction and indicate the potency of the training program to facilitate improvement in participants' rated competence and social skills. By tightening up the rater training component and examining students who are in greater need in training (i.e., students with lower baseline skill levels), it seems realistic to expect that the limited improvement of experimental subjects' ratings for this study points to a stronger showing of improvement in future studies of this type.

An interesting result of the analysis of the SSIT/skill and SSIT/anxiety ratings reveals an interaction effect between males in the experimental condition and males in the control condition. There was generally some increase in ratings of the posttest over pretest ratings for both groups of females and for experimental males; yet, control group males' rated skill and anxiety decreased. An explanation could be that the control group males perceived the exercise as a negative one, their only reward for participating in the study the receiving of class extra-credit





points. However, due to the small sample size of the experimental-group males (n=4), this result must be considered cautiously.

Instrument Validation

Disappointingly, the CSRS instrument revealed a lack of significant results in measuring the experimental group's improvement after attending the training program. On the CSRS, the raters were perceiving different things in the subjects' behavior. Nevertheless, the CSRS, like all other observation measures, is based on raters' subjective evaluations, and raters apply different standards because they are individuals. However, the beauty of the CSRS assessment is that it fills an assessment gap by providing specific information that can be diagnosed in almost any interpersonal context; it can be used as a self- or other- reference; and it can be used by trained or untrained raters. The caveat, however, is that the procedures for using it must be defined critically. In fact, Spitzberg (1986a) warns "if 'objective' information is desired, rater training and further scaling refinement is likely to be necessary for the raters to provide consistent information."

Regardless of its lack of significant results, the CSRS pretest scores did show a positive correlation with the SSIT/skill posttest scores — ratings made three weeks later -- as well as with the SSIT/anxiety pretest ratings. This correlation is important because it reveals that the CSRS is doing what it was intended to do: identifying competent behavior. As it should be, the CSRS is also positively correlated to itself, pretest to posttest. Heeding Spitzberg's caveat about rater training could indeed enable CSRS assessment to be a powerful instrument in measuring behavior comparisons such as this study.

## Implications

Methodologically speaking, for study designs like this one to provide more sensitive results, more thorough and highly structured rater training should be emphasized. More raters are also needed to facilitate higher reliability of the assessment instruments. If the time element is constrained (as it was for this study), using only one measure is recommended.

The training program schedule itself could be modified. As presented, the training program took place in two 6-hour sessions three weeks apart. Dissipation of results of the first session could have occurred in the interim.

The Interpersonal Skills Training Program could easily fill the time requirements for an entire 3-hour university course; perhaps more solid results could be obtained in that way. The training program as it was conducted in this study was designed and produced by graduate students in fulfillment of a course assignment. Up to now and including the training program that was measured in this study, there has been no real standard or continuity of design for presenting the training program since a different group of graduate students individualized it each semester (within the framework of the course guidelines). Now that the ISTP has been approved by the university, the training program's standard protocol should solidify and become more consistent. Future studies such as this one could possibly find stronger and more positive results for program participants.

## Summary

This study shows mixed results by instrument and by rater. However, taken as a whole, the data suggest considerable improvement for training





program participants' rated competence. The Interpersonal Skills Training Program offers significant potential for the enhancement of students' skills, as deduced from the improvement of training program participants' skill ratings on the Simulated Social Interaction Test and, according to one rater's scores, the Conversational Skills Rating Scale.

It is our opinion that better rater training and stricter selection criteria for subject participation (i.e., low competence levels) would facilitate significant, uniform improvement of participants' rated interpersonal skills.

REFERENCES


Bellack, A. (1979). A critical appraisal of strategies for assessing social skill. Behavioral Assessment, 1, 157-176.

Bellack, A. (1983). Recurrent problems in the behavioral assessment of social skill. Behavior Research and Therapy, 21, 29-41.

Curran, J. P. (1977). Skills training as an approach to the treatment of heterosexual-social anxiety: A review. Psychological Bulletin, 84, 140-157.

Curran, J. P. (1979a). Pandora's box reopened? The assessment of social skills. Journal of Behavioral Assessment, 1, 55-71.

Curran, J. P. (1979b). Social skills: Methodological issues and future directions. In A. S. Bellack & M. Herson (Eds.), Research and practice in social skills training (pp. 319-354). New York: Plenum.

Curran, J. P. (1982). A procedure for the assessment of social skills: The simulated social interaction test. In J. P. Curran and P. M. Monti (Eds.), Social skill training: A practical handbook for assessment and treatment (pp. 348-373). New York: Guilford Press..

Eisler, R. M. (1976). Behavioral assessment of social skills. In M. Hersen and A. S. Bellack (Eds.), Behavioral assessment: A practical handbook (pp. 369-395). Oxford, England: Pergamon.

Ellis, K., and Whittington, D. (1981). A guide to social skills training. London: Croon Helm.

Galassi, M. D. and Galassi, J. P. (1976). The effects of role playing variations on the assessment of assertive behavior. Behavior Therapy, 7, 343-347.

Hersen, M., and Bellack, A. S. (1976). Social skills training for chronic psychiatric patients: Rationale, research findings, and future directions. Comprehensive Psychiatry, 17, 559-580.

Hersen, M., and Bellack, A. S. (1977). Assessment of social skills. In A. R. Ciminero, K. S. Calhoun, & H. S. Adams (Eds.), Handbook of behavioral assessment (pp. 509-544). New York: John Wiley and Sons.







Horowitz, L. M., Inouye, D., and Siegelman, E. Y. (1979). On averaging judges' ratings to increase their correlation with an external criterion. Journal of Consulting & Clinical Psychology, 47, 453-458.

Kurtz, P. D., Marshall, E. K., and Banspach, S. W. (1985). Interpersonal skill-training research: A 12-year review and analysis. Counseling Education & Supervision, 24, 249-263.

Liberman, R. P. (1982). Assessment of social skills. Schizophrenia Bulletin, 8, 62-83.

McFall, R. M. (1982). A review and reformulation of the concept of social skills. Behavioral Assessment, 4, 1-33.

Spitzberg, B. H. (1985). Development and validation of a conversational skills rating scale. Unpublished manuscript.

Spitzberg, B. H. (1986a). A critical review of communication competence measures I: Self-reference instruments. Paper presented at SCA Conference, Chicago.

Spitzberg, B. H. (1986b). A critical review of communication competence measures II: Other-reference instruments. Paper presented at SCA Conference, Chicago.

Spitzberg, B. H. (in press). Measurement of communicative competence. In C. H. Tardy (Ed.), Instrumentation in communication research. New Jersey: Ablex.

Spitzberg, B. H., and Cupach, W. R. (1984). Interpersonal communication competence. Beverly Hills, CA: Sage.

Spitzberg, B. H., and Hurt, H. T. (1987). The measurement of interpersonal skills in instructional contexts. Communication Education, 36.

Strahan, R. F. (1980). More on averaging judges' ratings: Determining the most reliable composite. Journal of Consulting and Clinical Psychology, 48, 587-589.

Waltz, S. E., and Gough, H. G. (1984). External evaluation of efficacy by means of an adjective check list scale for observers. Journal of Personality and Social Psychology., 46, 697-704.

Table 1

Interrater Correlation Coefficient Assessed by Simple Pearson Correlation

| Variable | Coefficient | Significance |
|---|---|---|
| CSRS | | |
| Pretest | .53 | .01 |
| Posttest | .54 | .01 |
| SSIT/Skill | | |
| Pretest | .53 | .001 |
| Posttest | .56 | .001 |
| SSIT/Anxiety | | |
| Pretest | .53 | .001 |
| Posttest | .58 | .001 |

Table 2

Coefficient Alpha for the SSIT by Rater and Time

| SSIT Dimension | Rater 1 | Rater 2 |
|---|---|---|
| Skill | | |
| Pretest | .79 | .67 |
| Posttest | .68 | .81 |
| Anxiety | | |
| Pretest | .79 | .65 |
| Posttest | .73 | .82 |



Table 3

Correlated T-Tests Comparing Pre- and Posttest Competence Ratings in the Experimental and Control Groups

| | Experimental | | Control | |
|---|---|---|---|---|
| | Pre | Post | Pre | Post |
| Sample Means | 51.69 | 53.88 | 50.13 | 50.05 |
| S.D. | 3.62 | 5.59 | 5.33 | 6.64 |
| S.E. | 1.28 | 1.98 | 1.38 | 1.71 |
| t Value | | -1.57 | | 0.08 |
| p | | .16 | | .94 |
| n | | 8 | | 15 |

Table 4

Correlated T-Tests Comparing Pre- and Posttest Competence Ratings by Raters

| | Rater 1 | | Rater 2 | |
|---|---|---|---|---|
| | Pre | Post | Pre | Post |
| Means | 50.38 | 51.38 | 51.80 | 56.60 |
| S.D. | 3.19 | 5.66 | 6.61 | 7.39 |
| S.E. | 1.13 | 1.20 | 2.09 | 2.34 |
| t Value | | -0.41 | | -3.52 |
| p | | .689 | | .006 |
| n | | 8 | | 10 |

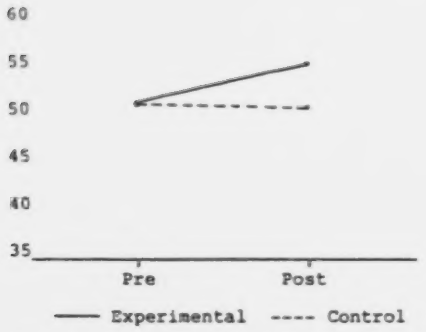


Fig. 1--A comparison of CSRS mean ratings of female subjects.

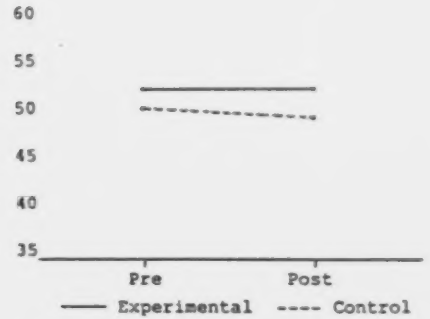


Fig. 2--A comparison of CSRS mean ratings of male subjects.



Table 5

Correlated T-Tests Comparing Pre- and Posttest SSIT Skill Ratings in the Experimental and Control Groups

| | Experimental | | Control | |
|---|---|---|---|---|
| | Pre | Post | Pre | Post |
| Sample Means | 42.59 | 54.41 | 47.05 | 47.68 |
| S.D. | 10.09 | 7.52 | 8.92 | 11.46 |
| S.E. | 3.04 | 2.27 | 1.99 | 2.56 |
| t Value | | -3.93 | | -0.32 |
| p | | .003 | | .752 |
| n | | 11 | | 20 |

Table 6

Correlated T-Tests Comparing Pre- and Posttest SSIT Anxiety Ratings in the Experimental and Control Groups

| | Experimental | | Control | |
|---|---|---|---|---|
| | Pre | Post | Pre | Post |
| Sample Means | 44.09 | 52.55 | 45.18 | 45.97 |
| S.D. | 9.01 | 4.66 | 7.99 | 9.87 |
| S.E. | 2.72 | 1.40 | 1.83 | 2.26 |
| t Value | | -3.49 | | -0.54 |
| p | | .006 | | .595 |
| n | | 11 | | 19 |

Table 7

Analysis of Variance of SSIT Skill
Posttest Ratings Crossed by
Condition and Sex

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Main Effects | 763.334 | 2 | 381.667 | 4.52 | .02 |
| Condition | 196.667 | 1 | 196.667 | 2.33 | ns |
| Sex | 375.456 | 1 | 375.456 | 4.45 | .04 |
| Interactions | 345.455 | 1 | 345.455 | 4.10 | .05 |
| Explained | 1108.790 | 3 | 369.597 | 4.38 | .01 |

$r^2 = .22$
$r = .46$

Table 8

Analysis of Variance of SSIT Anxiety
Posttest Ratings Crossed by
Condition and Sex

| | Sum of Squares | df | Mean Square | F | Sig. |
|---|---|---|---|---|---|
| Main Effects | 503.641 | 2 | 251.820 | 3.75 | .04 |
| Condition | 190.924 | 1 | 190.924 | 2.84 | ns |
| Sex | 166.779 | 1 | 166.779 | 2.48 | ns |
| Interactions | 39.431 | 1 | 39.431 | 0.59 | ns |
| Explained | 543.071 | 3 | 181.024 | 2.70 | ns |

$r^2 = .21$
$r = .46$



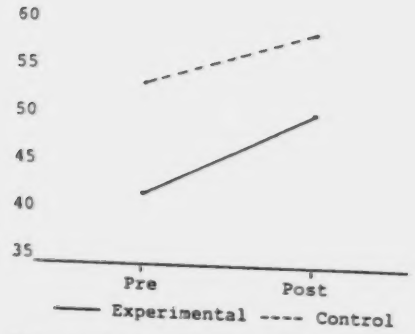


Fig. 3--A comparison of SSIT skill mean ratings of female subjects.

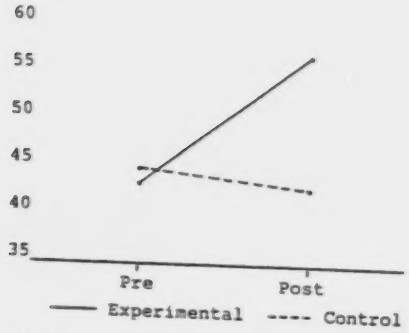


Fig. 4--A comparison of SSIT skill mean ratings of male subjects.

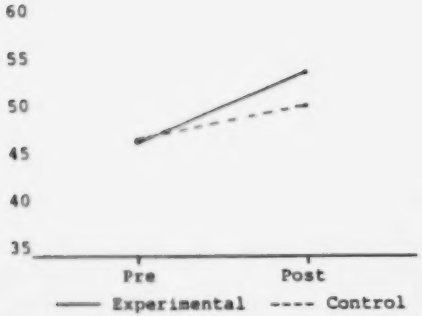


Fig. 5--A comparison of SSIT anxiety mean ratings of female subjects.

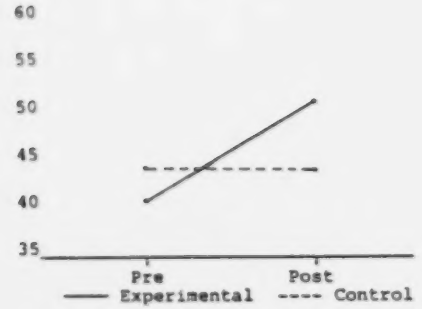


Fig. 6--A comparison of SSIT anxiety mean ratings of male subjects.